

DIALOG(R)File 347:JAPIO  
(c) 2007 JPO & JAPIO. All rts. reserv.

03268596      \*\*Image available\*\*  
VOICE RECOGNIZING DEVICE

PUB. NO.:        02-244096    [\*JP 2244096\*    A]  
PUBLISHED:      September 28, 1990 (19900928)  
INVENTOR(s):    KANEKO KAZUMA  
                 SHITAYA MITSUO  
APPLICANT(s):   MITSUBISHI ELECTRIC CORP [000601] (A Japanese Company or  
                 Corporation), JP (Japan)  
APPL. NO.:      01-066613    [JP 8966613]  
FILED:          March 16, 1989 (19890316)  
INTL CLASS:     [5] G10L-003/00; G10L-003/02  
JAPIO CLASS:    42.5 (ELECTRONICS -- Equipment)  
JAPIO KEYWORD: R108 (INFORMATION PROCESSING -- Speech Recognition &  
                 Synthesis)  
JOURNAL:        Section: P, Section No. 1144, Vol. 14, No. 569, Pg. 9,  
                 December 18, 1990 (19901218)

#### ABSTRACT

PURPOSE: To attain a high recognition rate without receiving the influence of ambient noises by generating noises by a noise generating means at the time of registration and comparing a standard voice characteristic pattern with a previously recorded input voice characteristic pattern in the same environment as the environment at the time of recognition, and thereby recognizing the above-mentioned pattern.

CONSTITUTION: A white noise signal generating section 3A generates a white noise electric signal N by receiving the instruction signal CN from a control section 2A when a switch 1 for registration is pushed. A speaker 4 converts the signal N to the white noise and outputs the same. The white noise is thus superposed on the voice of a specific speaker. This voice is passed through a microphone 5 and an analyzing section 6 and is thereby made into the standard voice characteristic pattern which is stored in a memory 7 for the standard pattern. A recognition mode is automatically started after the end of the registration. The voice inputted from the microphone 5 in the state of not operating the white noise signal generating section 3A and the speaker 4 is set through the analyzing section 6 and a memory 8 for the input pattern to a recognition section 9 where the voice is compared with the standard voice characteristic pattern and the inputted voice pattern is recognized. The influence of the ambient noises is eliminated in this way and the high voice recognition rate is attained.

## ⑫ 公開特許公報(A)

平2-244096

⑤ Int. Cl.<sup>9</sup>

G 10 L 3/00

3/02

識別記号

3 0 1 C

3 0 1

庁内整理番号

8842-5D

8842-5D

8842-5D

⑬ 公開 平成2年(1990)9月28日

審査請求 未請求 請求項の数 1 (全8頁)

⑭ 発明の名称 音声認識装置

⑯ 特 願 平1-66613

⑰ 出 願 平1(1989)3月16日

⑱ 発 明 者 金 子 和 磨 兵庫県尼崎市塚口本町8丁目1番1号 三菱電機株式会社  
産業システム研究所内⑲ 発 明 者 下 谷 光 生 兵庫県尼崎市塚口本町8丁目1番1号 三菱電機株式会社  
産業システム研究所内

⑳ 出 願 人 三菱電機株式会社 東京都千代田区丸の内2丁目2番3号

㉑ 代 理 人 弁理士 大岩 増雄 外2名

## 明 細 書

## 1. 発明の名称

音声認識装置

## 2. 特許請求の範囲

話者の音声に基づく入力音声特徴パターンと予め登録された前記話者の音声に基づく標準音声特徴パターンとを比較することにより入力音声パターンを認識する音声認識装置において、前記標準音声特徴パターンの登録時に雑音を前記話者に向けて発生する雑音発生手段を備えた事を特徴とする音声認識装置。

## 3. 発明の詳細な説明

(産業上の利用分野)

本発明は対雑音性能を向上させた音声認識装置に関する。

(従来の技術)

第8図は例えば特公昭61-42280号公報等の開示された従来の特定話者音声認識装置の主要部分の構成を示すブロック図である。同図において、1は登録用スイッチ、2はコントロール部、

5は音声を電気信号Sに変換するマイクロホン、6はマイクロホン5からの電気信号Sを分析して入力音声特徴パターンPを出力する分析部、7は登録時の入力音声特徴パターンPを標準音声特徴パターンとして記憶する標準パターン用メモリ、8は認識時の入力音声特徴パターンPを入力音声特徴パターンとして記憶する入力パターン用メモリ、9は入力音声特徴パターンと標準音声特徴パターンを比較して入力音声特徴パターンを認識した結果Rを出力する認識部である。なお、コントロール部2は登録用スイッチ1の状態に応じて符号6～9で示した要素の制御を音声認識用に行う。

従来の特定話者音声認識装置は上記のように構成され、登録モードと認識モードの2種類の動作モードを有する。

まず、登録モードの動作について説明する。登録用スイッチ1が押されるとON信号を発生する。このON信号を入力したコントロール部2は登録動作を開始する。マイクロホン5は音声を電気信号Sに変換する。分析部6はこの電気信号Sを入

力し、分析してその特徴を抽出したデジタルの入力音声特徴パターンPを出力する。標準パターン用メモリ7はコントロール部2の指示により分析部6から出力される入力音声特徴パターンPを標準音声特徴パターンとして記憶する。予め定められた単語数の標準音声特徴パターンを記憶することにより登録が終了する。この登録終了後、認識モードに自動的に移行する。

次に、認識モードでの動作について説明する。音声は、マイクロホン5により電気信号Sに音響-電気信号変換され、分析部6により分析されてその電気信号Sの特徴が抽出されたデジタルの入力音声特徴パターンPとなって分析部6から出力される。この入力音声特徴パターンPはコントロール部2の指示により入力パターン用メモリ8に記憶される。認識部9は、コントロール部2の指示により入力パターン用メモリ8に記憶された入力音声特徴パターンと標準パターン用メモリ7に先に記憶された標準音声特徴パターンとを比較演算し、標準音声特徴パターンの中から入力音声特

徴パターンに最も近似したパターンの単語を選び出し、この単語のパターンを認識結果Rとして出力する。

(発明が解決しようとする課題)

従来の音声認識装置は以上のように構成されているので、認識モードでの音声の入力時に周囲雑音が発録モード時の周囲雑音より大きい状況においては、特定話者の発声とその周囲雑音を原因とする心理的影響によって、登録モード時の発声に比べて著るしく変化したり、または、その周囲雑音が音声に重畳して入力される。これにより入力音声特徴パターンが、本来、近似すべき標準音声特徴パターンと異なったものになり、認識部9で認識できず、音声の認識率が低下する等の課題があった。

本発明は上記のような課題を解決するためになされたもので、周囲雑音の影響を受ける事なく、高い認識率を達成できる音声認識装置を得る事を目的とする。

(課題を解決するための手段)

本発明に係る音声認識装置は、標準音声特徴パターンの登録時に雑音を話者に向けて発生する雑音発生手段を設け、入力音声特徴パターンと標準音声特徴パターンとを比較して入力音声パターンを認識するようにしたものである。

(作用)

本発明における音声認識装置は、登録時に雑音発生手段により雑音を発生させ、認識時に周囲雑音が発生したと同じ環境下で標準音声特徴パターンを予め登録するためにそのパターンが入力音声特徴パターンに近似出来、認識可能となる。

(実施例)

以下、本発明の各実施例を図について説明する。第1図は本発明の一実施例に係る音声認識装置の構成を示し、同図において、従来装置と同一又は相当部分には第8図と同じ符号1、5～9を付しその説明を省略する。3Aは登録モード時にコントロール部2Aから指示信号C<sub>1</sub>を受けて白色雑音電気信号Nを発生する白色雑音信号発生部、4は白色雑音信号発生部3Aから白色雑音電気信号

Nを入力して白色雑音を発生するスピーカである。コントロール部2Aは、符号1、6～9の要素に接続され、従来装置のコントロール部2と同じ機能を有する他に、白色雑音信号発生部3Aに指示信号C<sub>1</sub>を出力する機能を有している。

第2図は第1図中の白色雑音信号発生部3Aの詳細な構成を示したブロック図である。第2図において、30は白色雑音データとしてのコードデータD<sub>1</sub>を予め記憶している白色雑音データ用読出し専用メモリ(以下、白色雑音データ用ROMと略称する。)、31は白色雑音データ用ROM30からのコードデータD<sub>1</sub>を適応差分パルス符号変調(ADPCM)してアナログの電気信号A<sub>1</sub>に変換する適応差分パルス符号変調用集積回路(以下、ADPCM用ICと略称する。)、32はADPCM用IC31の後段のローパスフィルタ(以下、LPFと略称する。)、33はLPF32の後段のスピーカ用アンプで、スピーカ4に接続されている。また、ADPCM用IC31はコントロール部2Aにも接続されている。

次に、第1図及び第2図を参照して第1実施例の動作について説明する。

まず、登録モードでの動作について説明する。登録用スイッチ1が押されるとコントロール部2Aは登録動作を開始する。白色雑音信号発生部3Aはコントロール部2Aからの指示信号C<sub>1</sub>を受けて白色雑音電気信号Nを発生する。この白色雑音信号発生部3Aの詳細な動作は以下の通りである。

ADPCM用IC31は、コントロール部2Aからの指示信号C<sub>1</sub>によって、白色雑音データ用ROM30に記憶されているコードデータD<sub>1</sub>を取出して適応差分パルス符号変調し、更にアナログ電気信号A<sub>1</sub>に変換して出力する。このアナログ電気信号A<sub>1</sub>はLPF32を通過してその低域が抽出された信号L<sub>1</sub>となり、更にスピーカ用アンプ33で増幅されて白色雑音電気信号Nに変換される。この時、コントロール部2Aは白色雑音データ用ROM30に記憶されているデータの再生時間長T<sub>1</sub>毎に指示信号C<sub>1</sub>をADPCM用IC

31に間歇的に送り続ける。従って、登録モードの間、白色雑音電気信号Nは白色雑音信号発生部3Aから連続して出力される。

スピーカ4は上記白色雑音電気信号Nを白色雑音に変換して特定話者に向けて出力する。特定話者の音声はスピーカ4からの白色雑音が重畳してマイクロホン5により音響-電気信号変換される。このマイクロホン5から出力される電気信号S<sub>1</sub>は分析部6を経由して標準音声特徴パターン(入力音声特徴パターンP<sub>1</sub>)化して標準パターン用メモリ7に記憶される。予め定められた単語数の標準音声特徴パターンが標準パターン用メモリ7に記憶されることによって登録は終了し、スピーカ4からの白色雑音の発生も終了する。登録終了後、自動的に認識モードに移行する。

次に認識モードでの動作について説明する。認識モード時にはコントロール部2Aから白色雑音信号発生部3Aに指示信号C<sub>1</sub>が出力されない。従って、白色雑音信号発生部3A及びスピーカ4は動作しない。この状態でマイクロホン5は音声

を電気信号S<sub>1</sub>に変換する。この電気信号S<sub>1</sub>は分析部6、入力パターン用メモリ8、認識部9により従来装置と同じ処理を受ける。これにより、認識部9から認識結果R<sub>1</sub>が出力される。

本実施例の場合、登録モード時における標準音声特徴パターンは、白色雑音が聞こえている状況時の特定話者の発声に白色雑音が重畳した音をパターン化している。また、認識モード時における入力音声特徴パターンは、白色雑音の大きさ程度の周囲雑音が聞こえている状況時の特定話者の発声に周囲雑音が大なり小なり重畳した音をパターン化している。従って、同じ単語に対する標準音声特徴パターンと入力音声特徴パターンとが極めて近似し、認識部9により認識可能となる。

なお、上記実施例では、ADPCM用ICを備えたものを示したが、ADPCM用ICの代わりにデジタル/アナログ(D/A)変換器やPCM用IC等の他の再生用ICを用いても上記実施例と同様の効果を奏する。

第3図は本発明に係る音声認識装置の他の一実

施例を示し、装置全体の構成は第1図に示した装置と同じであるが、白色雑音信号発生部3Aの構成が第1実施例と異なる。第3図において、34は白色雑音が磁気テープに記録された白色雑音用カセットテープ、35は白色雑音用カセットテープ34を装着し、コントロール部2Aから指示信号C<sub>1</sub>を受けて再生するテープデッキ、33はテープデッキ35から出力される再生信号を増幅して白色雑音電気信号Nにしてスピーカ4に出力するスピーカ用アンプである。登録モード時にテープデッキ35はコントロール部2Aからの指示信号C<sub>1</sub>を受けると白色雑音用カセットテープ34の再生を開始し、白色雑音用カセットテープ34に記録されている磁気信号B<sub>1</sub>を電気信号T<sub>1</sub>に変換する。スピーカ用アンプ33はこの電気信号T<sub>1</sub>を増幅して白色雑音電気信号Nにしてスピーカ4に出力する。

一方、認識モード時には、コントロール部2Aから指示信号C<sub>1</sub>が出力されないためにテープデッキ35やスピーカ4は動作しない。その他の動

作は第1実施例と同じなのでその説明を省略する。

なお、第2実施例ではテープデッキとカセットテープを使用したか、この代りにコンパクトディスク(CD)プレイヤーと白色雑音用CD-ROMを使用しても上記第2実施例と同様の効果を奏する。

また、第1及び第2実施例では白色雑音を発生するものも示したが、ROMやカセットテープにピンク雑音や自動車走行時の雑音等のデータを記録しておき、ピンク雑音や自動車走行時の雑音等を発生しても上記実施例と同様の効果を奏する。

第4図は本発明の第3実施例に係る装置全体の構成を示し、第4図において、第1実施例と同一部分には第1図と同符号1、4～9を付し、その説明を省略する。コントロール部2Bは、第1実施例のコントロール部2Aと同じ機能を有する他に、雑音選択用スイッチ10から雑音の種類を選択するための雑音選択信号を入力すると選択雑音信号発生部3Bに雑音選択用の指示信号C<sub>n</sub>を出力する機能を有する。この選択雑音信号発生部

ール部2Bからの指示信号C<sub>n</sub>によって指示された雑音の選択雑音信号を発生する。この選択雑音信号発生部3Bの詳細な動作は以下の通りである。

ADPCM用IC31は、コントロール部2Bからの指示信号C<sub>n</sub>の指示に従って、第1～第Jの雑音データ用ROM36-1…36-Jのいずれかを選択して雑音コードデータD<sub>n</sub>を読出して適応差分パルス符号変調し、更にアナログ電気信号A<sub>n</sub>に変換して出力する。このアナログ電気信号A<sub>n</sub>はLPF32を通過してその低域が抽出された信号L<sub>n</sub>となり、更にスピーカ用アンプ33で増幅されて選択雑音電気信号Nに変換される。この時、コントロール部2Bは、ADPCM用IC31にコードデータD<sub>n</sub>の再生時間長T<sub>n</sub>毎に指示信号C<sub>n</sub>を出力する。

この選択雑音信号Nはスピーカ4に供給されて雑音に変換される。

次に認識モードの場合の動作について説明する。認識モード時にはコントロール部2Bは指示信号C<sub>n</sub>を選択雑音信号発生部3Bに出力しない。従

3Bは指示信号C<sub>n</sub>の種類に応じた雑音信号を発生するもので、その詳細な構成を第5図に示す。また、スピーカ4は選択雑音信号発生部3Bの出力端子に接続されている。

第5図において、36-1, 36-2…36-Jは第1～第Jの雑音データ用ROMで、各々が互いに異なった種類の雑音データを予め記憶している。その他、ADPCM用IC31、LPF32、スピーカ用アンプ33の構成は第2図に示した構成と同じである。但し、ADPCM用IC31は、コントロール部2Bからの指示信号C<sub>n</sub>の種類に応じて第1～第Jの雑音データ用ROM36-1…36-Jのいずれかを選択してデータコードD<sub>n</sub>を読出す。

次に第4図及び第5図を参照して第3実施例の動作について説明する。登録モード時において、雑音選択用スイッチ10により雑音の種類が選択されると、この雑音選択信号を入力したコントロール部2Bは雑音の種類を指示する指示信号C<sub>n</sub>を出力する。選択雑音信号発生部3Bはコントロ

って、選択雑音信号発生部3B及びスピーカ4は動作せず、雑音を発生しない。

その他の音声認識の動作については第1実施例の動作と同じなのでその説明を省略する。

なお、1種類の雑音データを1つのROMに記憶させたが、複数種類の雑音データを1つのROMに記憶させておき、その雑音データのいずれかをアドレス指定により選択して読出しても良い。

第6図は本発明の第4実施例に係る装置全体の構成を示し、第6図において、第1実施例と同一部分には第1図と同符号1、4～9を付し、その説明を省略する。コントロール部2Cは第1実施例のコントロール部2Aと同じ機能を有する他に、録音用スイッチ11が押された時の録音モードか、再生モードかを指示する指示信号C<sub>n</sub>を雑音信号記録再生部3Cに出力する機能を有する。この雑音信号記録再生部3Cは、スピーカ4や雑音用マイクロホン12にも接続され、指示信号C<sub>n</sub>に応じて雑音の録音又は再生を行うもので、その詳細な構成を第7図に示す。

第7図において、37はバックアップ用バッテリー38によって電源を供給されているランダムアクセスメモリ（以下、RAMと略称する。）、31AはADPCM用ICで、コードデータD<sub>n</sub>をADPCM処理してアナログ電気信号A<sub>n</sub>に変換したり、逆にアナログ電気信号A<sub>n</sub>をコードデータD<sub>n</sub>に変換する。マイクロホン12、マイク用アンプ39、LPF32Aがこの順に接続され、LPF32、スピーカ用アンプ33、スピーカ4がこの順に接続されている。ADPCM用IC31AはRAM37、LPF32Aの出力端子、LPF32の入力端子に接続されていると共にコントロール部2Cにも接続されている。

次に第6図及び第7図を参照して第4実施例の動作について説明する。かかる構成の音声認識装置は登録モード、認識モードに加えて録音モードを有する。

まず、録音モードの動作について説明する。録音用スイッチ11が押されると、コントロール部2Cは録音モードを指示する指示信号C<sub>n</sub>を雑音

RAM37から先に格納したコードデータD<sub>n</sub>を読出し、ADPCM処理してアナログ電気信号A<sub>n</sub>に変換する。このアナログ電気信号A<sub>n</sub>はLPF32→スピーカ用アンプ33を経由してスピーカ4に供給される。これにより、スピーカ4は録音された雑音を出力する。

次に認識モード時の場合の動作について説明する。認識モード時にはコントロール部2Cから指示信号C<sub>n</sub>が出力されないために雑音信号録音再生部3C、スピーカ4が動作せず、録音や再生を行わない。

その他の音声認識の動作については第1実施例で述べた動作と同じなのでその説明を省略する。

上記第4実施例によれば、音声認識装置を使用する環境下の雑音を録音でき、その録音した雑音を登録モード時に発生させることができるので、任意の雑音環境下においても高い認識率を達成することができる。

なお、上記各実施例において、雑音を発生するスピーカの代りにヘッドホンを使用し、このヘッ

ダホンにより雑音を話者に聞かせても上記実施例と同様の効果を奏する。

雑音用マイクロホン12は周囲雑音を電気信号N<sub>n</sub>に変換する。この電気信号N<sub>n</sub>は、マイク用アンプ39により増幅されて電気信号L<sub>n</sub>にされ、更にLPF32Aを通過して低域のみのアナログ電気信号A<sub>n</sub>となる。このアナログ電気信号A<sub>n</sub>を入力したADPCM用ICは、コントロール部3Cから録音モードの指示信号C<sub>n</sub>を入力していれば、このアナログ電気信号A<sub>n</sub>を雑音データとしてのコードデータD<sub>n</sub>に変換してRAM37に格納する。所定の時間についてコードデータD<sub>n</sub>を記録すると録音モードは終了する。RAM37はバックアップ用バッテリー38からデータを保持するための必要な電圧Vの供給を受けているために格納されたデータを保持し続ける。

次に登録モード時の場合の動作について説明する。登録モード時にはそのモードを指示する指示信号C<sub>n</sub>を入力したADPCM用IC31Aは

ドホンにより雑音を話者に聞かせても上記実施例と同様の効果を奏する。

#### 〔発明の効果〕

以上のように、本発明によれば標準音声特徴バターンの登録時に話者に向けて雑音を発生させ、前記話者の入力音声パターンと予め登録された標準音声特徴パターンとを比較して入力音声パターンを認識するように構成したので、周囲雑音の影響を排除することができ、高い音声認識率を達成できる効果がある。

#### 4. 図面の簡単な説明

第1図は本発明の第1実施例に係る音声認識装置のブロック図、第2図は第1図中の白色雑音信号発生部の構成を示すブロック図、第3図は本発明の第2実施例における白色雑音信号発生部の構成を示すブロック図、第4図は本発明の第3実施例に係る音声認識装置のブロック図、第5図は第4図中の選択雑音信号発生部の構成を示すブロック図、第6図は本発明の第4実施例に係る音声認識装置のブロック図、第7図は第6図中の雑音信

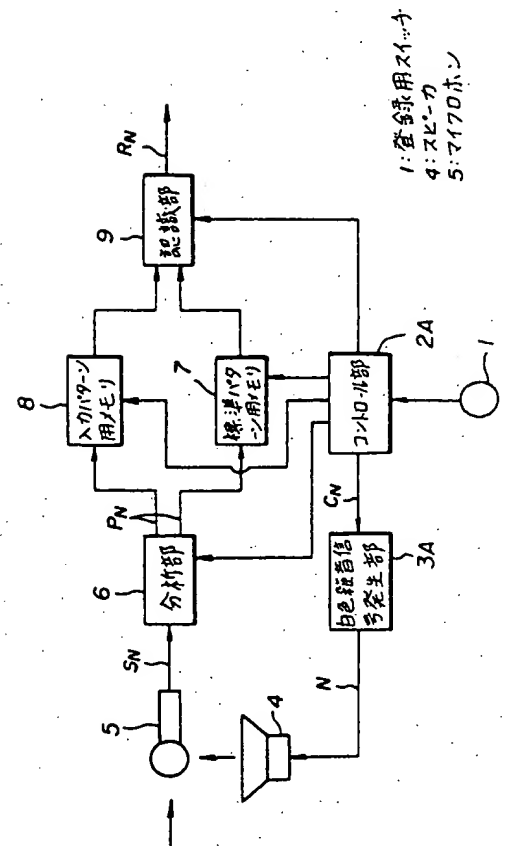
号記録再生部等の構成を示すブロック図、第8図は従来の音声認識装置のブロック図である。

図中、1…登録用スイッチ、2A、2B、2C…コントロール部、3A…白色雑音信号発生部、3B…選択雑音信号発生部、3C…雑音信号記録再生部、4…スピーカ、5…マイクロホン、6…分析部、7…標準パターン用メモリ、8…入力パターン用メモリ、9…認識部、10…雑音選択用スイッチ、11…録音用スイッチ、12…雑音用マイクロホン。

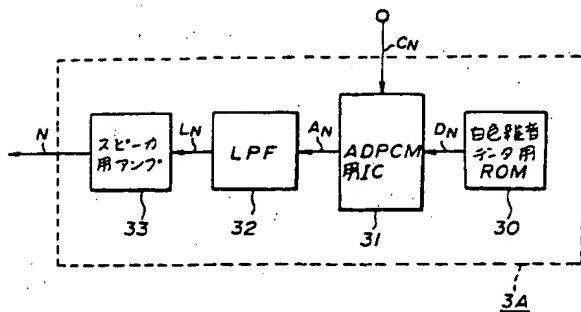
なお、図中同一符号は同一、又は相当部分を示す。

代理人 大 岩 増 雄

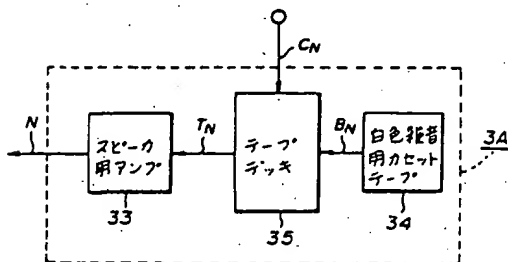
第 1 図



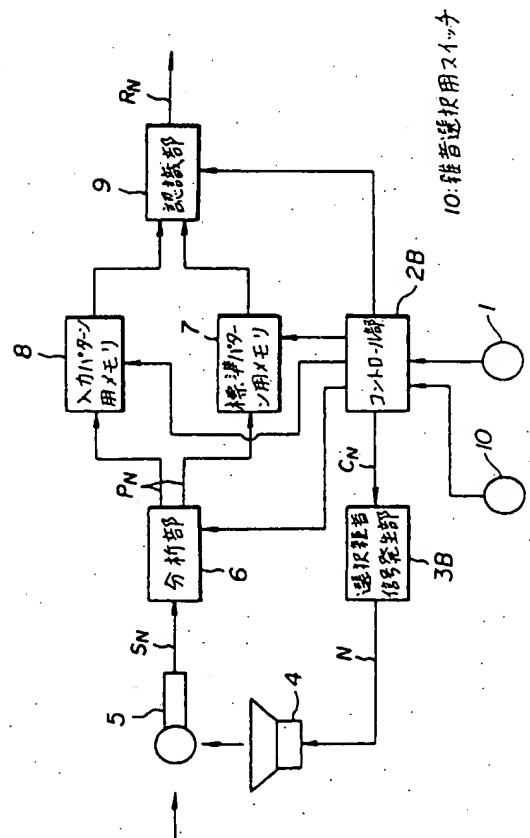
第 2 図



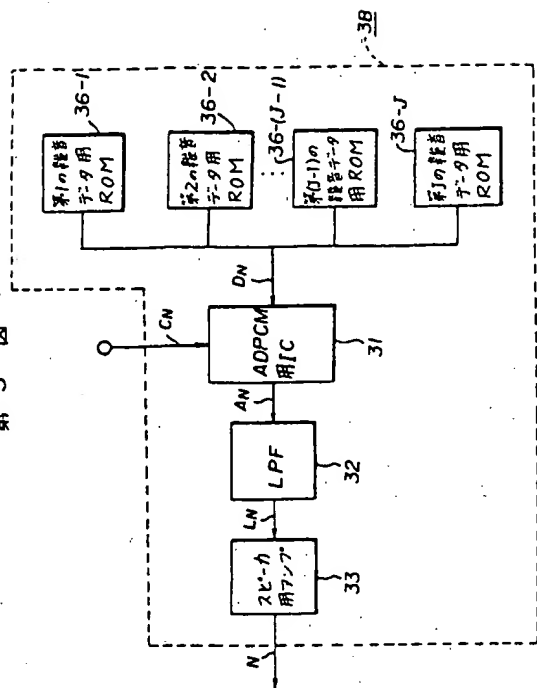
第 3 図



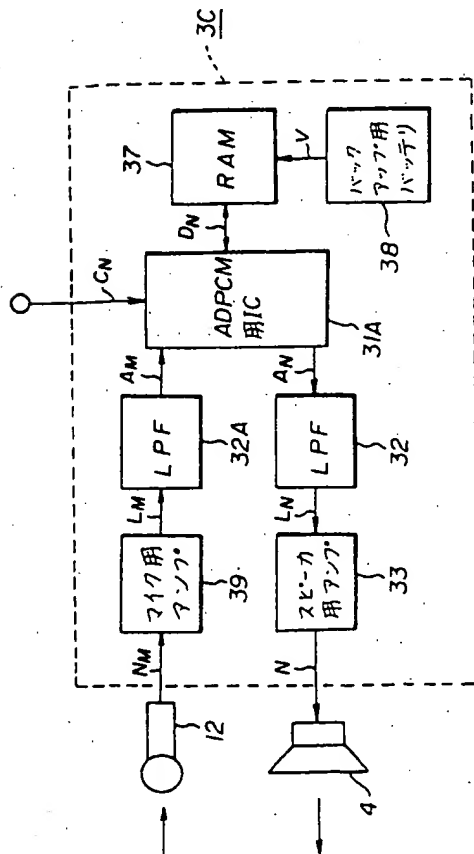
第 4 図



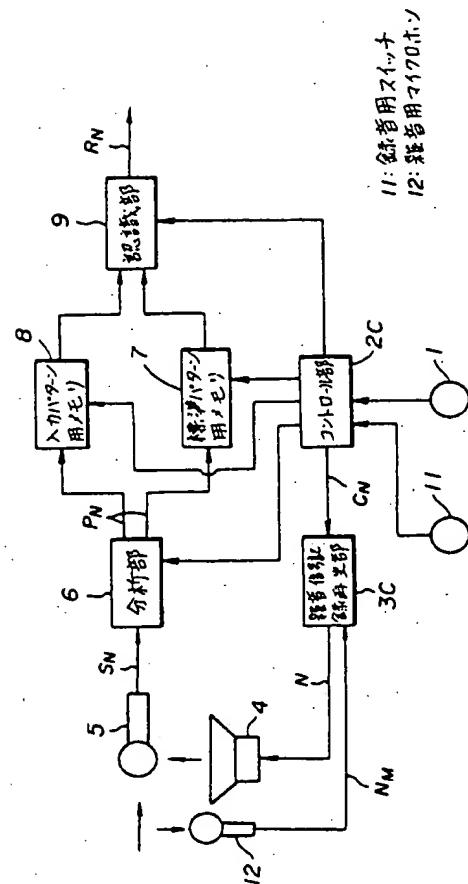
第 5 図



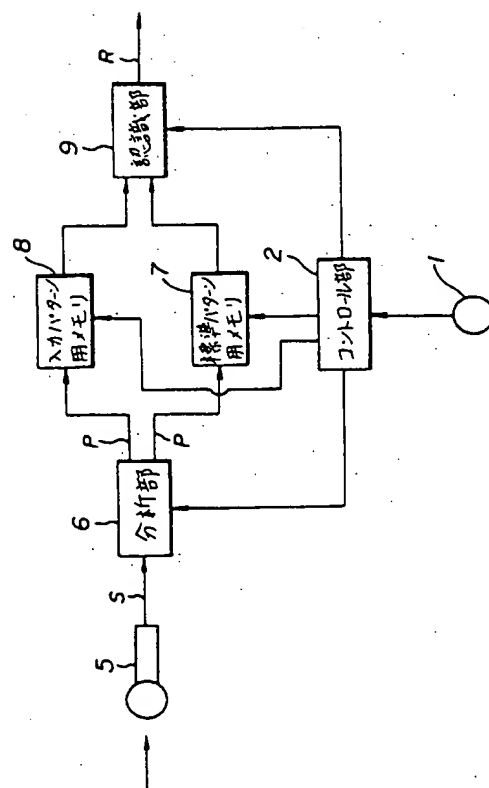
第 7 図



第 6 図



第 8 図



11: 録音用スピーカ  
12: 録音用マイクホン



手続補正審(自発)  
1 11 29  
平成 年 月 日  
適

特許庁長官殿

1. 事件の表示 平 特願 1-66613号

2. 発明の名称 音声認識装置

3. 補正をする者

事件との関係 特許出願人  
住 所 東京都千代田区丸の内二丁目2番3号  
名 称 (601)三菱電機株式会社  
代表者 志 岐 守 哉

4. 代 理 人

住 所 東京都千代田区丸の内二丁目2番3号  
三菱電機株式会社内  
氏 名 (7375)弁理士 大 岩 増 雄  
(連絡先03(213)3421特許部)

5. 補正の対象

明細書の「発明の詳細な説明」の欄

方式  
審査



6. 補正の内容

- (1) 明細書第4頁2行目の「単語のパターン」を「単語」と補正する。
- (2) 同第4頁17行目の「受ける事なく、」を「低減し、」と補正する。
- (3) 同第5頁8行目～同頁11行目の「発生手段により……となる。」を「発生手段が認識時の周囲雑音を模擬して雑音が発生させるので、登録音声パターンが認識時の入力音声パターンに近づき、認識性能が向上する。」と補正する。
- (4) 同第8頁5行目、同第8頁5行目～同頁6行目、同第9頁7行目、同第9頁10行目の「特定話者」を「話者」と補正する。
- (5) 同第16頁2行目の「記録」を「録音」と補正する。
- (6) 同第18頁9行目の「排除」を「低減」と補正する。

以 上